



*Escuela Técnica Superior de  
Ingenieros de Caminos, Canales  
y Puertos.*



UNIVERSIDAD DE CANTABRIA

# ANÁLISIS DE LOS EXTREMOS DE LLUVIA Y CAUDAL EN LA CUENCA DEL RÍO PAS: INUNDACIONES DE ENERO DE 2019

Trabajo realizado por:

*Ainara Manrique Díez*

Dirigido:

*Jesús Casado Rodríguez*

*Manuel del Jesus Peñil*

Titulación:

**Máster Universitario  
en Gestión Integrada  
de Sistemas Hídricos**

Santander, octubre de 2019

TRABAJO FINAL DE MASTER

## Índice

1	Introducción .....	3
1.1	Inundaciones de Cantabria enero 2019 .....	3
1.2	Objetivos .....	3
2	Área de estudio y datos .....	4
2.1	Cuenca del Pas .....	4
2.1.1	Situación .....	4
2.1.2	Características y morfología .....	4
2.1.3	Clima e hidrología .....	5
2.2	Obtención de los datos .....	7
2.3	Observaciones iniciales .....	7
3	Metodología .....	9
3.1	Tratamiento de los datos .....	9
3.1.1	Obtención de las series caudal .....	9
3.1.2	Obtención de las series de precipitación .....	10
3.2	Análisis de extremos .....	14
3.2.1	Análisis univariado .....	14
3.2.2	Análisis bivariados .....	15
4	Resultados .....	21
4.1	Análisis univariado .....	21
4.1.1	Valores extremos para periodos de retorno predefinidos .....	21
4.1.2	Periodos de retorno de la inundación de enero de 2019 .....	21
4.2	Análisis bivariados .....	22
4.2.1	Correlación entre variables .....	22
4.2.2	Cópulas gaussianas .....	24
5	Discusión .....	28
6	Propuestas .....	31
7	Referencias .....	32
8	Anexos .....	33
8.1	Datos de precipitación de las estaciones seleccionadas para la cuenca A088 entre 1950 y 2019 .....	33
8.2	Periodos de retorno para los valores de precipitación, nivel y caudal .....	34
8.3	Correlación entre los extremos .....	36
8.4	Ajuste de la distribución de extremos .....	37

## Índice de figuras

Figura 1 Comunidad Autónoma de Cantabria y situación de la cuenca de estudio .....	4
Figura 2 Situación de todas las estaciones de medición en las cercanías de la cuenca del Pas-Pisueña.....	6
Figura 3 Hietograma e hidrograma para los días de inundación usando estaciones de la cuenca del Pas y el aforo Q104 .....	7
Figura 4 Comparación de las series de caudal y nivel para los aforos A088 y 1215 .....	9
Figura 5 Serie temporal resultante tras la unión de los aforos 1215 y A088 .....	10
Figura 6 Polígonos de Thiessen y estaciones asociadas para la cuenca del río Pas .....	11
Figura 7 Ejemplo de mapa de precipitación diaria en la cuenca de estudio interpolado por IDW.....	13
Figura 8 Series de máximos anuales para la precipitación, nivel y caudal para las series obtenidas .....	15
Figura 9 Ejemplo de representación conjunta de las funciones de distribución de dos variables: probabilidad de ocurrencia de inundación y nivel máximo del río. ....	16
Figura 10 Ejemplo de representación de las funciones de distribución de dos variables: probabilidad de ocurrencia de inundación y nivel máximo del río, al aplicar una cópula. ....	17
Figura 11 Representación gráfica de los escenarios AND y OR para el cálculo del periodo de retorno bivariado .....	18
Figura 12 Ejemplo de cópula gaussiana con las curvas de periodo de retorno para los escenarios AND y OR .....	19
Figura 13 Funciones de distribución empíricas y analíticas para la precipitación, el caudal y el nivel .....	21
Figura 14 Precipitación máxima y sus principales correlaciones .....	23
Figura 15 Caudales máximos y sus principales correlaciones.....	23
Figura 16 Nivel máximo y sus principales correlaciones .....	24
Figura 17 Representación de la cópula entre las variables de precipitación y media de la precipitación de dos últimos días. ....	25
Figura 18 Representación de la cópula entre las variables de media de la precipitación de dos y tres últimos días, con sus probabilidades. ....	26

## Índice de tablas

Tabla 1 Características principales de la cuenca del Pas-Pisueña .....	5
Tabla 2 Estaciones AEMET en la cuenca de estudio y sus periodos de retorno obtenidos hasta 2010.....	6
Tabla 3 Valores de precipitación registrados en los pluviómetros de la cuenca para los días de inundación y previos.....	7
Tabla 4 Estaciones pluviométricas seleccionadas para realizar la serie completa de precipitación y su localización .....	12
Tabla 5 Ejemplo de valor anómalo en una serie de precipitación (en mm).....	12
Tabla 6 Ajuste de la función de distribución de extremos y valor extremo para diversos periodos de retorno para la precipitación, el caudal y el nivel .....	21
Tabla 7 Periodos de retorno de las principales variables de estudio para los días de inundación.....	22
Tabla 8 Correlación de Spearman entre las diferentes variables de estudio.....	22
Tabla 9 Tabla con los valores de los puntos de mayor probabilidad para cada periodo de retorno para las variables de precipitación y precipitación media de dos días .....	26
Tabla 10 Tabla con los valores de los puntos de mayor probabilidad para cada periodo de retorno para las variables de precipitación media de dos y tres días .....	27

# 1 Introducción

## 1.1 Inundaciones de Cantabria enero 2019

Las inundaciones en la comunidad autónoma de Cantabria son un problema recurrente que impacta de forma grave las infraestructuras, actividades económicas y vidas de sus habitantes.

Para gestionar estas emergencias, desde 2010 existe el Plan Especial de Protección Civil de la Comunidad Autónoma de Cantabria ante el Riesgo de Inundaciones, INUNCANT. El 24 de enero del 2019 se activó el Nivel 2 de alerta, que implica la movilización completa de toda la estructura del Plan, cambiando de la fase de prealerta que llevaba desde el 21 de enero, debido al temporal de nieve y lluvia.

Las inundaciones comenzaron el 23 de enero en varias localidades y hasta el día 26 de enero a las 13:30 no fue desactivado el nivel 2, afectando en ese tiempo a 61 municipios. Los servicios de emergencias atendieron 1686 llamadas, se desalojaron 156 personas, varias infraestructuras, como puentes y carreteras, quedaron destruidas o fuera de servicio durante días a causa de los desprendimientos de tierras y la inundación.

Un primer análisis del comportamiento de las lluvias durante la inundación no parece corresponderse con la magnitud del impacto en la región, llevando a sospechar que usar únicamente los valores de la precipitación para explicar la gravedad de este evento son insuficientes.

## 1.2 Objetivos

El principal objetivo de este informe es verificar que las inundaciones ocurridas en Cantabria a finales de enero del 2019, en concreto para la cuenca del río Pas, fue la conjunción de varias variables, y no solamente el valor de la precipitación, las que generaron las inundaciones en la cuenca.

Partiendo de esta idea, se identifican las otras variables que pudieron haber influido en la subida del nivel de las aguas. Se elaboran series de precipitación y caudal para realizar un análisis univariado para los periodos de retorno de las inundaciones de enero de 2019. Seguidamente, se realiza un análisis multivariable mediante cópulas gaussianas para identificar qué conjunto de variables tuvieron una mayor importancia en estas inundaciones.

## 2 Área de estudio y datos

### 2.1 Cuenca del Pas

#### 2.1.1 Situación

La Comunidad Autónoma de Cantabria se encuentra al norte de la Península Ibérica. Limitando al norte con el mar Cantábrico y al sur con la Cordillera Cantábrica, la cual se extiende de forma paralela a la costa y muy cerca de la misma, lo cual conforma un relieve muy accidentado en todo el territorio.

La cuenca del Pas-Pisueña se encuentra en la región central de Cantabria. Es una de las cuencas con mayor extensión, nace entre los montes de Valnera y de Samo, en la zona Vasco-Cantábrica de la Cordillera y limitando con la provincia Burgos.



Figura 1 Comunidad Autónoma de Cantabria y situación de la cuenca de estudio

#### 2.1.2 Características y morfología

La mayoría de los ríos del territorio presentan una orientación sur-norte, naciendo en las montañas y vertiendo sus aguas en el mar Cantábrico, debido a la poca distancia entre las montañas y la costa, suelen ser ríos cortos, de poca extensión y de topografía irregular.

La cuenca del Pas-Pisueña presenta unas características similares al resto de cuencas de Cantabria, con fuertes desniveles entre la Cordillera y la franja litoral. En las zonas altas está conformada por pequeños ríos y arroyos, de comportamiento muy torrencial, que se transforman en valles anchos en las zonas bajas y litoral.

La cuenca está compuesta por dos ríos principales: el Pas y el Pisueña, tiene una extensión superficial de 620 Km<sup>2</sup> y poseen una longitud de unos 60 Km. Sus principales afluentes son el arroyo Pandillo, el de la Magdalena y el río Yera. Ambos ríos se unen a la altura de Vargas y desembocan en la ría de Mogro.

Cuenca	Superficie (km <sup>2</sup> )	Perímetro (m)	Altitud media (m)	Pendiente media (%)	Tiempo concentración (h)
Pas- Pisueña	620,40	144,2	495	34,7	22,58

Tabla 1 Características principales de la cuenca del Pas-Pisueña

La cubierta vegetal está compuesta en las zonas altas por bosques de haya y robles, con algunas zonas de pastizales en las laderas dedicadas a la ganadería, en las zonas bajas de los valles existen bosques mixtos de frondosas, y bosques de ribera compuestos por alisos y sauces.

En las zonas más llanas, a medida que los ríos se adentran en la zona litoral, aumentan las concentraciones urbanas y la cantidad de prados dedicados a la agricultura y ganadería. La cuenca del Pas presenta una densidad de población de 52 hab./km<sup>2</sup>, mientras que la del Pisueña de 78 hab./km<sup>2</sup>.

### 2.1.3 Clima e hidrología

Cantabria presenta un clima templado atlántico, influido por el clima mediterráneo. Su situación entre los límites de las masas de aires polares y tropicales genera cambios bruscos de tiempo y una precipitación abundante a lo largo de todo el año. La región también muestra una estacionalidad bien definida en las lluvias, siendo los meses de mayor pluviosidad los comprendidos entre octubre y abril.

La precipitación media en Cantabria es de 1353 mm/año, distribuyéndose espacialmente de forma homogénea, aumentando desde el litoral a la Cordillera Cantábrica en bandas más o menos definidas, con algunas excepciones de carácter local.

Una de estas excepciones es el efecto barrera provocado por las montañas en el nacimiento del Pas, las más altas de la zona. Este efecto provoca que las nubes se mantengan sobre las cumbres, haciendo que sea una de las zonas con mayor cantidad de lluvias recogidas, más de 2400 mm/año, mientras que en áreas como las cuencas del Soba y de la Magdalena, no se superen los 900 mm/año.

Las frecuentes precipitaciones y la extensión reducida de la cuenca conllevan unos tiempos de concentración cortos que dificultan la previsión de eventos de inundación. Estos eventos han ocurrido con más frecuencia en los márgenes del río Pas, entre las localidades de Vega del Pas y Puente Viesgo, y en el curso bajo desde Puente Viesgo a su desembocadura. Hasta 2010, el número de inundaciones históricas registradas en la cuenca era de 17 para el Pas (el máximo en toda Cantabria) y 5 para el Pisueña.



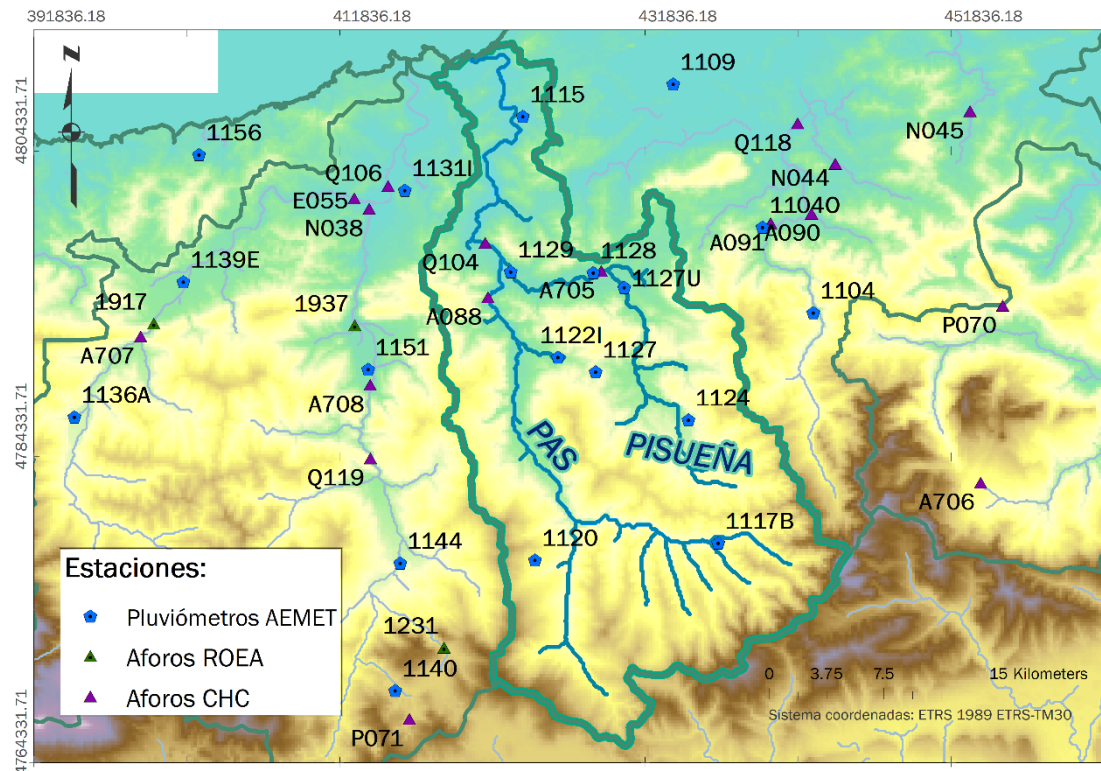


Figura 2 Situación de todas las estaciones de medición en las cercanías de la cuenca del Pas-Pisueña

Cod. Estación AEMET	Municipio	Altitud (m)	Precipitación media diaria (mm)	Precipitación máxima diaria (mm)	Periodo de retorno Precipitación máxima diaria (mm)		
					T 50	T100	T 500
1115	Piélagos	13	3,7	166,4	154	186	288
1116	San Pedro del Romeral	780	4	130	109	117	132
1117	Vega de Pas	380	5,2	180	167	188	243
1120	Luenta	537	4,2	200	132	150	200

Tabla 2 Estaciones AEMET en la cuenca de estudio y sus periodos de retorno obtenidos hasta 2010

## 2.2 Obtención de los datos

Las series de precipitación se obtuvieron de los pluviómetros de la AEMET y del SAIH-CHC (Sistema Automático de Información Hidrológica de la Confederación Hidrográfica del Cantábrico). Las series de caudal y nivel se obtuvieron de las estaciones del ROEA (Red Oficial de Estaciones de Aforos) y del SAIH-CHC. Para la delimitación de la cuenca se utilizó el modelo digital del terreno de 25 m de resolución disponible en el Centro Nacional de Información Cartográfica.

## 2.3 Observaciones iniciales

Las precipitaciones se adentraron en la región de oeste a este, comenzando las inundaciones en algunas zonas el día 23, sin embargo, el valor máximo de las lluvias recogido en la región del Pas-Pisueña para ese día es de 86.3 mm en un solo pluviómetro.

	Estaciones			
Fecha	Q104	A705	A708	9012
2019-01-23	47.7 mm	54.8 mm	86.3 mm	73.1 mm
2019-01-24	67.8 mm	65.6 mm	119.4 mm	84.3 mm

Tabla 3 Valores de precipitación registrados en los pluviómetros de la cuenca para los días de inundación y previos

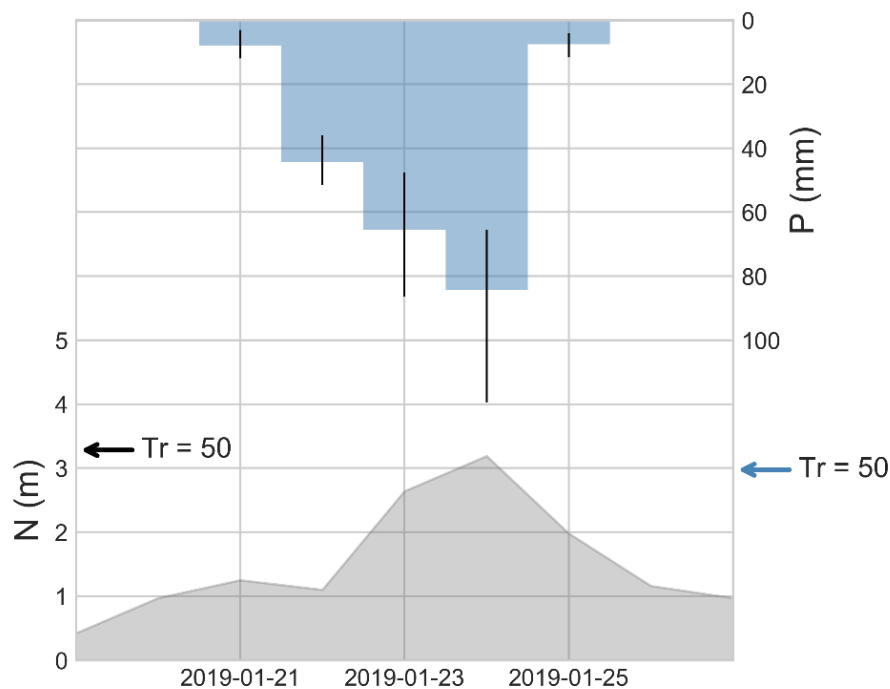


Figura 3 Hietograma e hidrograma para los días de inundación usando estaciones de la cuenca del Pas y el aforo Q104



En la imagen se ve en la parte superior el hietograma para cuatro pluviómetros situados en la cercanía del cauce del Pas-Pisueña, reflejando su media y la variación entre valores máximos y mínimos. En la parte inferior se muestra el hidrograma de aforo Q104, que recoge las aguas tanto del Pas como del Pisueña. En los lados de la figura pueden verse marcados con flechas los períodos de retorno, para la precipitación (derecha) son los obtenidos hasta 2010 según datos históricos, para el nivel (izquierda) han sido calculados para los años 2010-2019, el periodo para el que se disponen datos en este aforo.

Se observa que, si bien el nivel alcanzado por el río tiene un periodo de retorno de 50 años, la precipitación diaria observada no llega a los 140 mm correspondientes a dicho retorno de 50 años. Esto lleva a considerar que quizá el valor diario de la precipitación no es la única variable que influyó en el desbordamiento de los ríos.

Entre los años 60 y 82 del siglo XX se realizaron diversas actuaciones en esta cuenca para evitar inundaciones como las que habían ocurrido años atrás. Pese a las medidas tomadas, tanto en 2016 como en 2019 se han producido nuevas avenidas. Por lo tanto, es importante analizar si la forma de dimensionar infraestructuras y otras actuaciones de gestión de inundaciones, basadas en el periodo de retorno de una única variable, como la precipitación diaria, es realista.

## 3 Metodología

### 3.1 Tratamiento de los datos

Para poder realizar el análisis estadístico de las inundaciones ocurridas, necesitamos conocer el comportamiento hidrológico de la cuenca y de los fenómenos extremos que hayan tenido lugar en ella, para poder compararlos con lo ocurrido en las recientes inundaciones. Por tanto, el primer paso es reunir series temporales lo suficientemente largas como para que resulten representativas de la cuenca.

#### 3.1.1 Obtención de las series caudal

En la cuenca existen tres aforos en funcionamiento actualmente: Q104 (Vargas), A088 (Puente Viesgo) y A705. De estos, únicamente el aforo A088 tiene valores de caudal y nivel para los días de las inundaciones.

Las series temporales de estos aforos resultan limitadas, empezando entre 2008 y 2011, por lo que apenas se tendrían 11 años para realizar los análisis. Afortunadamente, el aforo 1215, de la ROEA, contiene datos de caudal y nivel desde 1969 y 2015 y está situado en el mismo punto del río Pas que el aforo A088, por lo que se decide unir ambos aforos para obtener una serie larga y representativa del río.

##### 3.1.1.1 Unión de los caudales en A088.

Las series solapan entre los años 2008 y 2015. En esos años se observan algunos días con discrepancias entre las medidas de ambos aforos, pero en la mayoría de los casos son diferencias pequeñas, así que se considera que tienen el suficiente parecido como para poder unir ambas series para conseguir una serie larga más adecuada para nuestros cálculos.

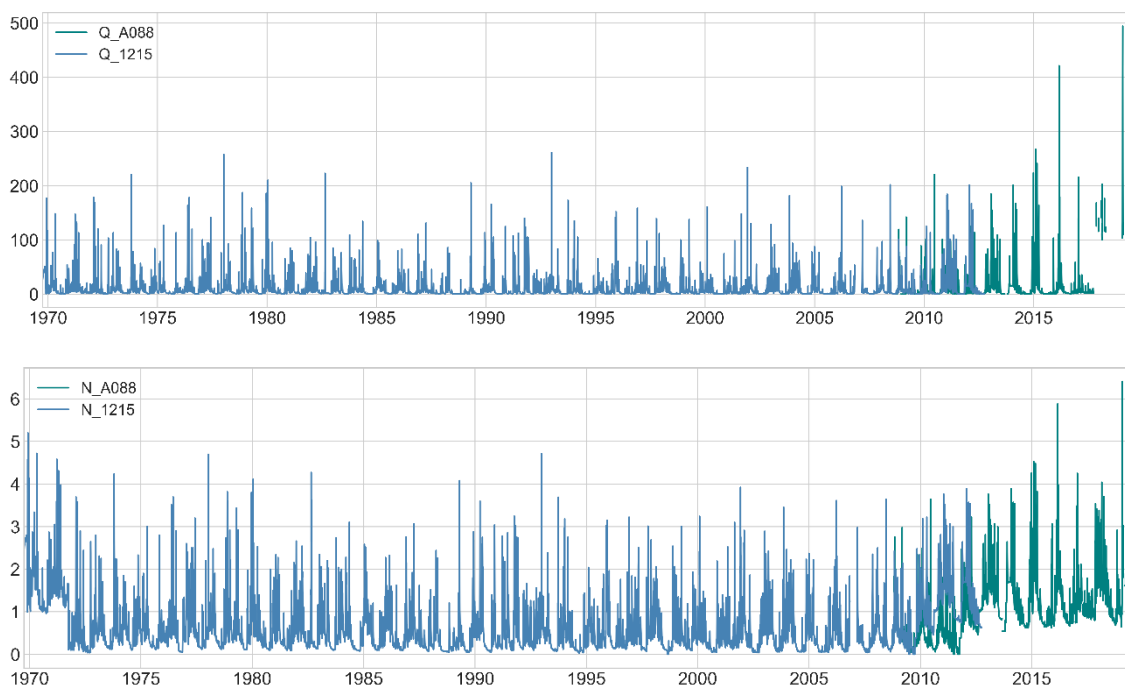


Figura 4 Comparación de las series de caudal y nivel para los aforos A088 y 1215

A la hora de decidir qué datos emplear en las fechas donde los datos no coinciden, se opta por el aforo A088 de la Confederación, al mostrar una mayor correlación con los valores de precipitación.

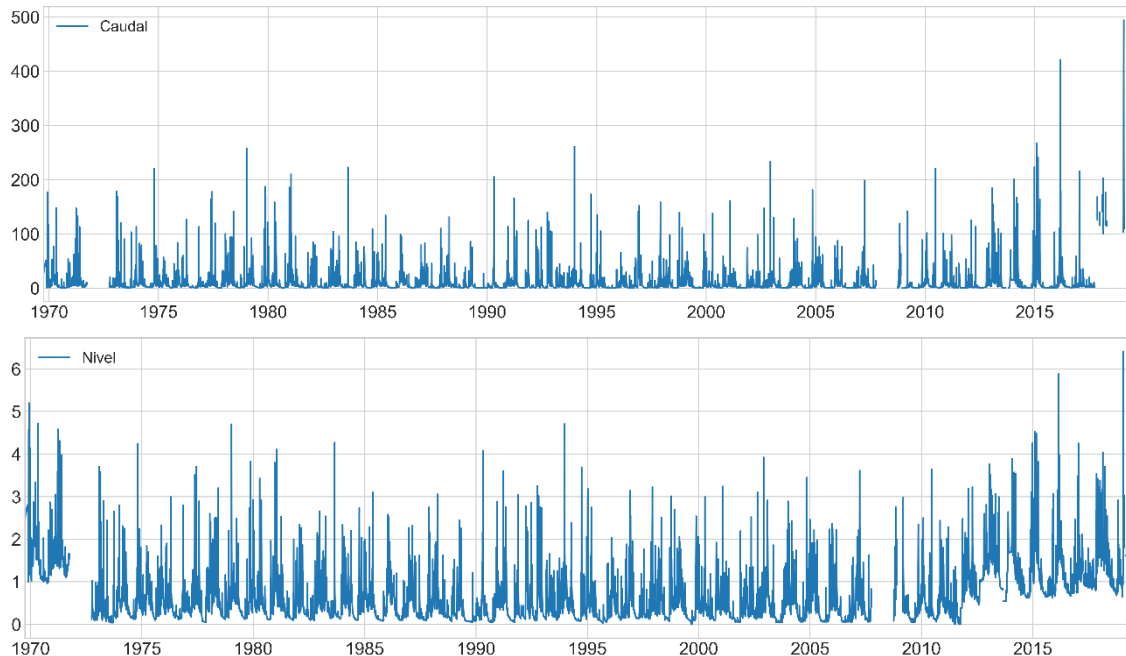


Figura 5 Serie temporal resultante tras la unión de los aforos 1215 y A088

### 3.1.2 Obtención de las series de precipitación

#### 3.1.2.1 Selección de estaciones y aforos

En una primera selección, para conocer qué estaciones tienen mayor influencia sobre nuestra cuenca de estudio, se emplearán los polígonos de Thiessen. Las áreas de los polígonos están trazadas de forma que el perímetro que dibuja marca la equidistancia entre las diferentes estaciones. Bajo este método, los polígonos de cada estación representan su área de influencia sobre la cuenca de estudio.

Los aforos A088 y 1215 (Puente Viesgo) se encuentra aguas arriba de la confluencia de los ríos Pas y Pisueña, por lo que el área de estudio se reduce a la de cierre en el aforo A088 que solo recoge la cuenca del río Pas.

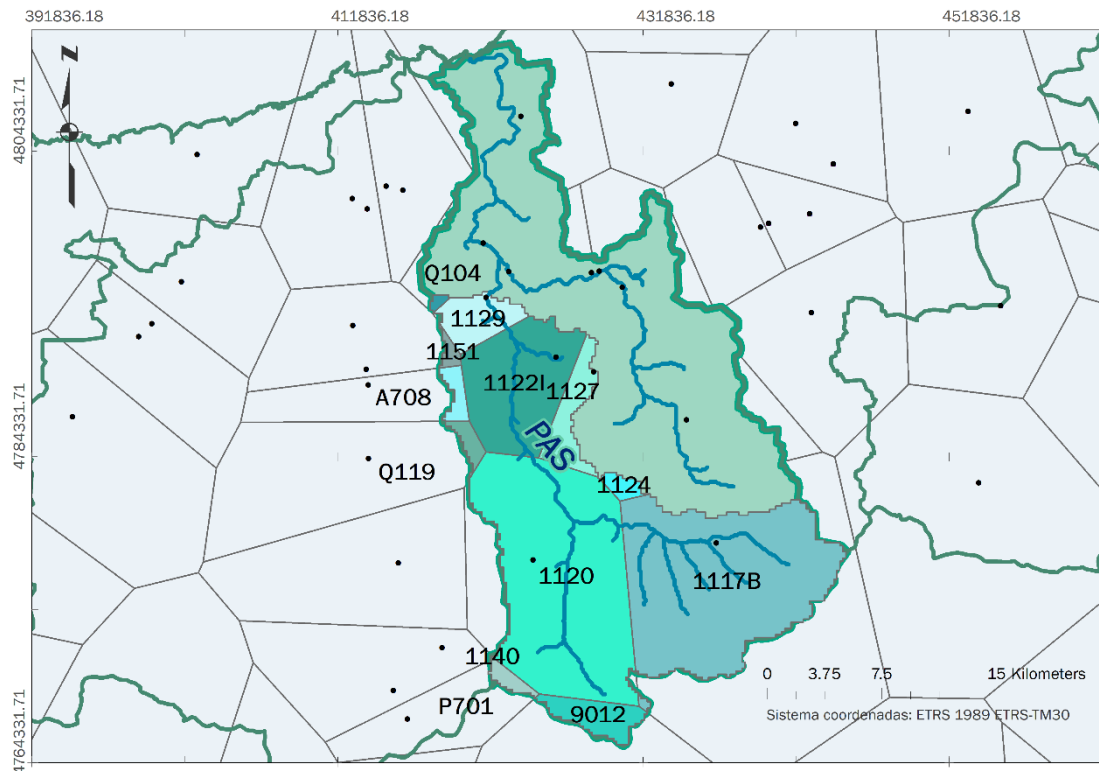


Figura 6 Polígonos de Thiessen y estaciones asociadas para la cuenca del río Pas

### 3.1.2.2 Estaciones seleccionadas

Las estaciones seleccionadas para el cálculo de la serie de precipitación en la cuenca serían, finalmente, las siguientes:

CÓDIGO	COORD_X	COORD_Y
P071	416409.04	4767180.46
1120	424630.00	4777578.00
1140	415481.00	4769044.00
1122I	426124.00	4790834.00
A708	413840.29	4789023.36
Q104	421371.00	4798299.00
1151	413712.00	4790048.00
1117B	436611.00	4778689.00
Q119	413866.00	4784214.00
1127	428598.00	4789875.00

<b>9012</b>	422557.00	4760386.00
<b>1129</b>	423031.00	4796422.00
<b>1124</b>	434656.00	4786737.00

Tabla 4 Estaciones pluviométricas seleccionadas para realizar la serie completa de precipitación y su localización

### 3.1.2.3 Análisis inicial de los datos

Tras la primera selección de las estaciones se analizarán los datos que contienen para asegurarnos de que son adecuados para nuestro estudio.

#### 3.1.2.3.1 Anomalías

En la serie de precipitación de AEMET existen una serie de códigos que señalan valores erróneos, se encontraron -3 y -4, que representan precipitación inapreciable y precipitación acumulada respectivamente. En el primer caso se decidió redondear el valor a 0, en el segundo se eliminó el dato ya que no se dispone del valor real.

Algunas series también presentan datos que son claramente anómalos, como valores de lluvia muy por encima de lo razonable en la cuenca o cuyos valores no guardan relación alguna con los datos obtenidos en estaciones cercanas.

	1122I	1120	1151	1124	1127
28/02/1984	1.9	0	3.6	7	2.3
29/02/1984	5.1	200	4.4	9	7.5
01/03/1984	8.3	0	0.3	0	0
02/03/1984	5.3	0	2	12	1.5

Tabla 5 Ejemplo de valor anómalo en una serie de precipitación (en mm)

Para evitar que estos datos anómalos puedan falsear los cálculos, se han considerado como nulos.

#### 3.1.2.3.2 Datos nulos

El conjunto de las series seleccionadas para los datos de lluvia presenta valores desde 1950 hasta 2019. Aunque en la mayoría de ellas hay huecos de varios años, se ha considerado que hay suficientes días recogidos como para que los huecos no representen un problema para representar el comportamiento de la precipitación de la cuenca.

#### 3.1.2.4 Cálculos de las series de precipitación histórica

Para realizar los análisis de este estudio fue necesario obtener una única serie de precipitación entre 1950 y 2019, usando las series de las estaciones seleccionadas anteriormente, para ello se usó el método la distancia inversa ponderada.

#### 3.1.2.5 Distancia inversa ponderada

El método de la distancia inversa ponderada consiste en interpolar el valor de la precipitación en un punto como la media ponderada de la precipitación en las

estaciones cercanas. La ponderación se hace a través del inverso de la distancia (elevado a un exponente), de manera que las estaciones más cercanas al punto de interpolación tienen un mayor peso.

$$P = \sum_{i=1}^n w_i P_i = \sum_{i=1}^n \frac{d_i^{-p}}{\sum_{i=1}^n d_i^{-p}} P_i \quad 1$$

Donde  $P$  es el valor de la precipitación a obtener,  $w_i$  el peso de las distancias para cada estación  $i$ ,  $n$  el número de estaciones,  $d$  la distancia de las estaciones al punto a calcular y  $p$  la potencia, que sirve para valorar la importancia de la distancia en el cálculo, para este análisis se ha optado por un valor de 2.

Los cálculos se realizarán mediante Python para poder aplicar el cálculo a cada día para una serie larga de precipitación sobre un mapa en formato ASCII que proporciona los valores del área y las distancias. El cálculo se hará empleando el algoritmo K-nearest neighbor (vecinos más cercanos), que realizará la labor de interpolar los valores en cada punto del área introducida. Este método permite así valorar la influencia de la distancia en los cálculos, incluida la altura, aunque para esta cuenca se observó que incluir las alturas apenas aportaba diferencia en los resultados, por lo que finalmente no se incluyeron.

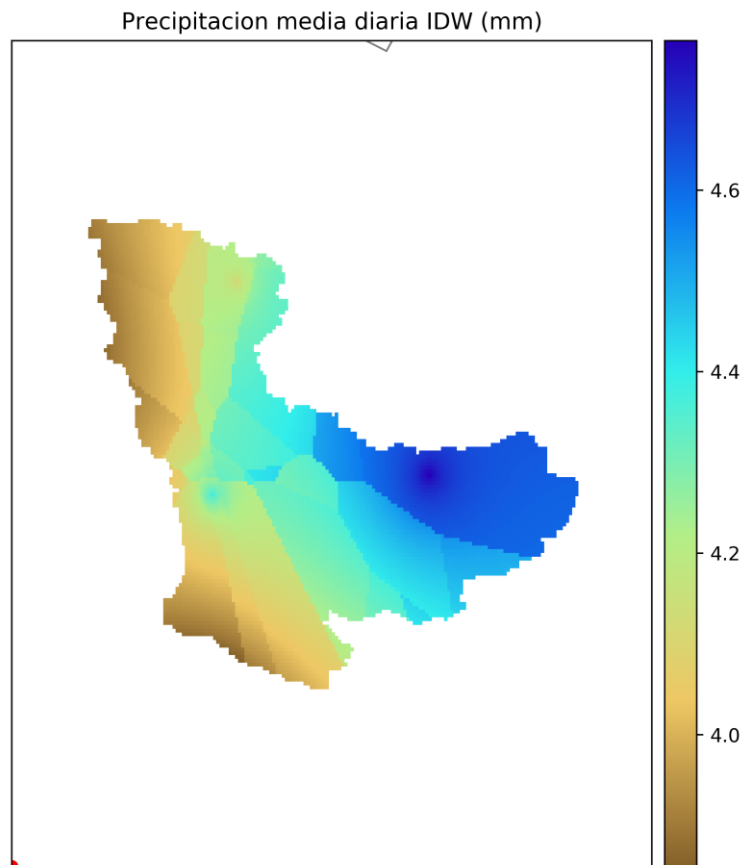


Figura 7 Ejemplo de mapa de precipitación diaria en la cuenca de estudio interpolado por IDW

Como resultado obtendremos una única serie de valores de precipitación diarios para toda el área de la cuenca.



### 3.1.2.6 Otras variables

Las variables que se han considerado usar para realizar los análisis bivariados, además de las ya mencionadas series de lluvia, caudal y nivel, son los siguientes:

- Precipitación del día previo (1d\_P\_mm).
- Intensidad de lluvia media de los últimos 2 días (2d\_P\_mm).
- Intensidad de lluvia media de los últimos 3 días (3d\_P\_mm).
- Intensidad de lluvia media de los últimos 6 días (6d\_P\_mm).

Las series de precipitación acumulada en los días previos son una aproximación a la saturación del suelo, que como es bien sabido afecta directamente sobre la generación de escorrentía.

## 3.2 Análisis de extremos

### 3.2.1 Análisis univariado

#### 3.2.1.1 Periodos de retorno

Para realizar una primera aproximación a la excepcionalidad de las inundaciones de enero de 2019, se hará un análisis univariado de frecuencia de extremos, calculando los periodos de retorno de las variables de precipitación diaria, caudal y nivel. Para el análisis de frecuencia de extremos se generó la serie de máximos anuales de cada una de las variables de estudio (Figura 8).

El periodo de retorno define la recurrencia en años de un evento extremo. Se calcula como la inversa de la probabilidad de excedencia (ecuación 2).  $F_X$  es la función de distribución que define la probabilidad de no excedencia, es decir, la probabilidad de que una variable,  $X$ , no supere un determinado valor  $x$ .

$$T_X = \frac{1}{1 - F_X(x)} = \frac{1}{1 - P(X < x)} \quad 2$$

La función de distribución empleada habitualmente para el análisis de extremos en hidrología es la GEV (generalized extreme values), que engloba tres funciones de distribución de extremos: Gumbel, Fréchet y Weibull.

$$F_X = \exp \left[ - \left\{ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right\}^{-\frac{1}{\xi}} \right] \quad 3$$

La ecuación 3 muestra la función GEV definida por tres parámetros: localización ( $\mu$ ), forma ( $\xi$ ) y escala ( $\sigma$ ) de la distribución. El parámetro de forma ( $\xi$ ) controla el comportamiento de la distribución y define si la función es una Gumbel ( $\xi = 0$ ), Fréchet ( $\xi > 0$ ) o Weibull ( $\xi < 0$ ).

Los cálculos mediante la GEV se realizan en base a los máximos anuales obtenidos en las series.

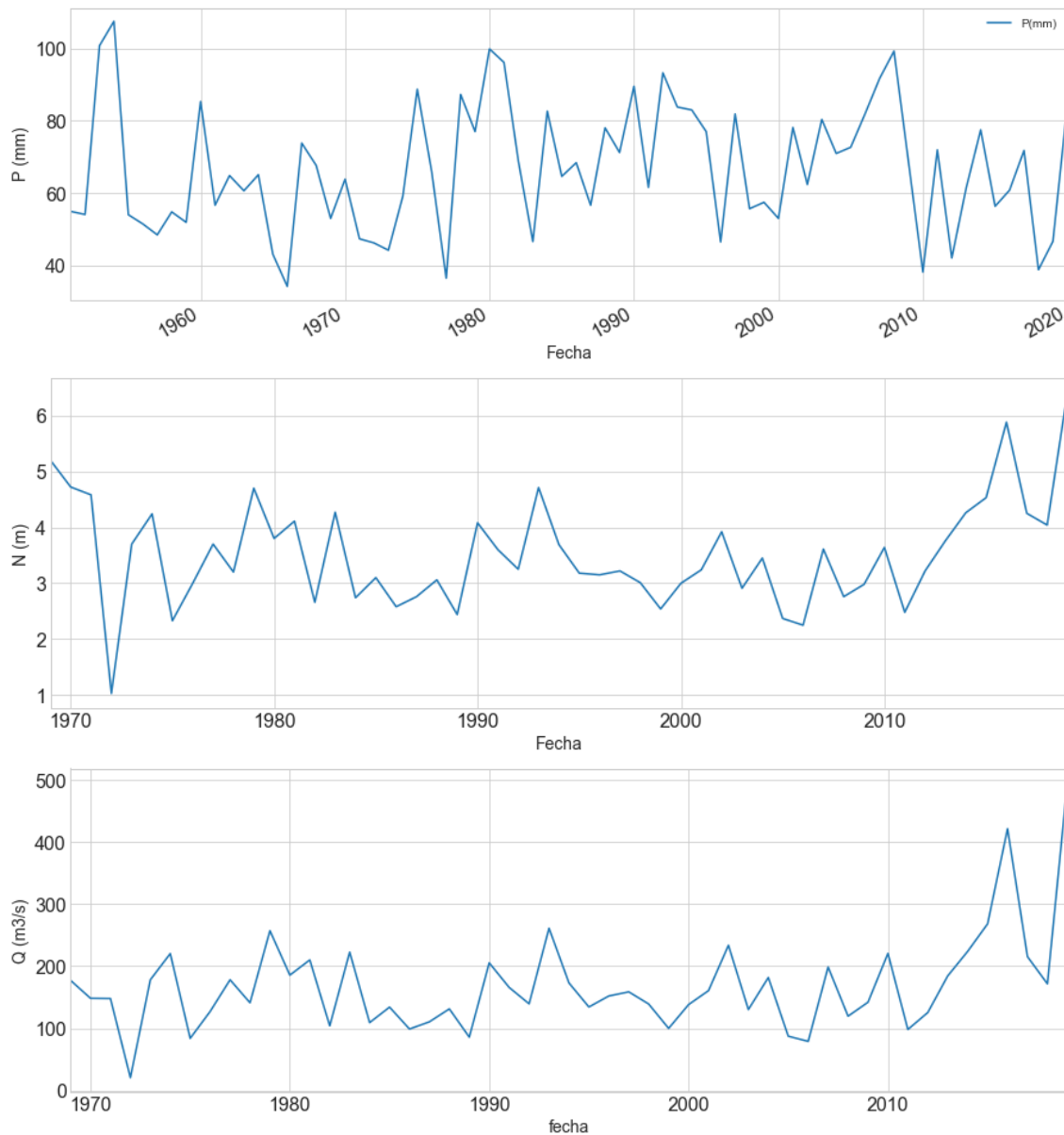


Figura 8 Series de máximos anuales para la precipitación, nivel y caudal para las series obtenidas

### 3.2.2 Análisis bivariados

El primer paso antes de realizar las cópulas sería conocer la correlación que existen entre las variables que vamos a analizar, para poder centrar los últimos cálculos en aquellas variables que muestren mejores relaciones.

La correlación indica la fuerza y dirección entre dos variables, calcularemos primero las series temporales mediante el método de Spearman, ya que este método se calcula comparando la jerarquía de las variables, que resulta más relevante para el este análisis que la comparación de datos en bruto.

Para definir aún más las variables a escoger se realizarán correlaciones partiendo de los máximos anuales de nivel, caudal y precipitación diaria. Se calcularán por separado, ya que los máximos no siempre coinciden en los mismos días.

### 3.2.2.1 Cópulas gaussianas

Las cópulas son una función de distribución multivariada que nos permiten conocer la relación entre dos variables con valores aleatorios, uniendo sus funciones de distribución multivariada mediante su distribución marginal.

Al tener dos variables X e Y, el cálculo de sus probabilidades de ocurrencia conjunta para obtener los periodos de retorno, se calculan habitualmente uniendo sus funciones de distribución. Generando un modelo de probabilidad como el de la figura 9.

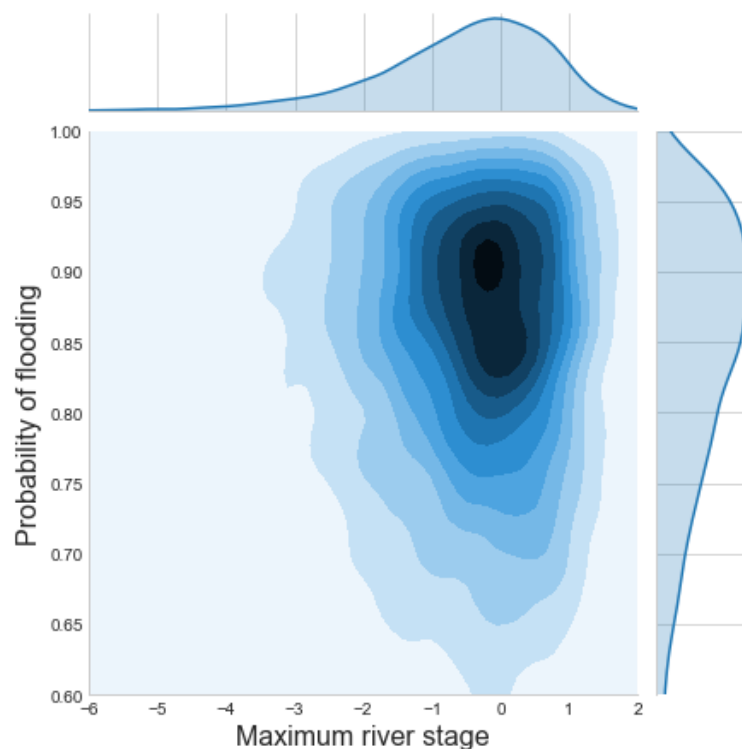


Figura 9 Ejemplo de representación conjunta de las funciones de distribución de dos variables: probabilidad de ocurrencia de inundación y nivel máximo del río.

Estas técnicas dan por hecho que las distribuciones de las variables provienen de la misma familia y que su dependencia sigue una distribución lineal, lo cual no es necesariamente el caso en cálculos hidrológicos.

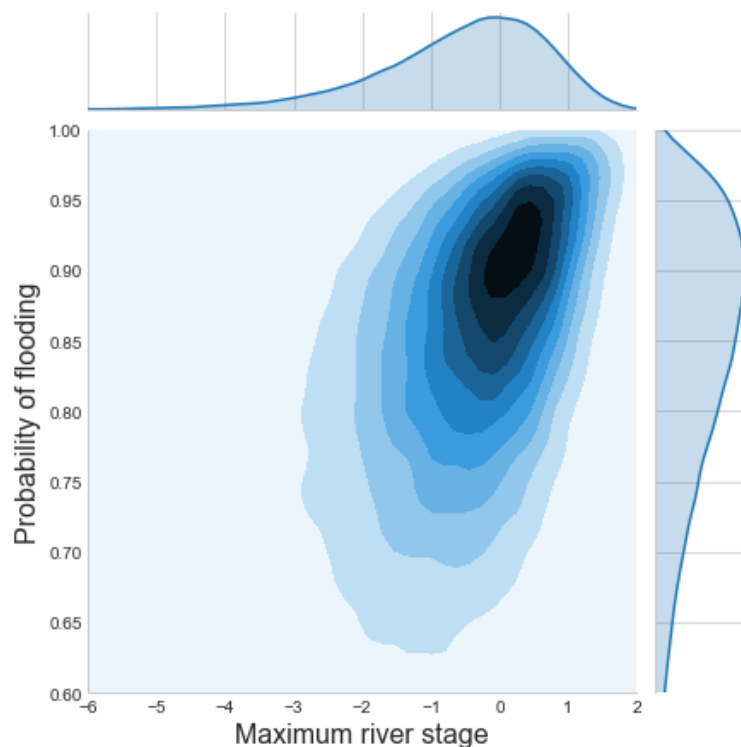
Las cópulas están basadas en el teorema de Sklar, que para el caso de dos variables que nos ocupa, puede definirse como que la función de distribución de cualquier par de variables aleatorias continuas  $F_{XY}(x,y)$  puede definirse como:

$$F_{XY}(x,y) = C\{F(x), G(y)\} = C(u,v), x,y \in \mathfrak{R} \quad 4$$

Donde  $F(x)$ , representada como  $u$ , y  $G(y)$ , representada como  $v$ , son las distribuciones marginales de las variables  $X$  e  $Y$ , y la función  $C: [0,1]^2 \rightarrow [0,1]$  es la cópula.

Esto quiere decir que, a diferencia de los modelos de distribución conjunta, las cópulas nos permiten calcular la distribución de la marginal en función del comportamiento conjunto de las variables a analizar, permitiendo modelar la estructura de dependencia de las variables sin necesidad de conocer la familia de distribuciones a las que pertenecen, al ser  $U$  y  $V$  (variables de la función marginal) univariantes.

Como resultado, aplicando una cópula al ejemplo anterior, obtenemos una representación de modelo de probabilidad más definido.



*Figura 10 Ejemplo de representación de las funciones de distribución de dos variables: probabilidad de ocurrencia de inundación y nivel máximo del río, al aplicar una cópula.*

A efectos prácticos, los pasos empleados para el cálculo de las cópulas en este análisis serán:

- 1) Ajuste de las marginales
- 2) Generar la cópula gaussiana bivariada a partir de las medias y la matriz de correlación entre las dos variables.
- 3) Hacer el análisis de extremo sobre la cópula gaussiana (ambas marginales son una  $N[0,1]$ )
- 4) Transformar los resultados de la cópula gaussiana para reproducir las marginales del paso 1

### 3.2.2.2 Escenarios

Para el cálculo de los periodos de retorno bivariados partimos de dos escenarios:

El escenario  $OR^V$  indica la probabilidad de que una de las variables,  $X$  o  $Y$ , supere un determinado umbral  $x$  o  $y$ .

$$T^V(u, v) = \frac{1}{1 - C(u, v)} \quad 5$$

Donde  $T^V$  representa el periodo de retorno de la cópula  $C(u, v)$  para el escenario  $OR^V$ . La cópula corresponde a la probabilidad de no excedencia de ambos umbrales, por lo que la ecuación sigue una forma similar al caso univariado.

El escenario  $AND^A$  indica la probabilidad de que ambas variables superen simultáneamente los umbrales  $(x, y)$ .

$$T^A(u, v) = \frac{1}{1 - u - v + C(u, v)} \quad 6$$

Donde  $T^A$  representa el periodo de retorno para el escenario  $AND^A$ , obtenido a partir de la probabilidad de no excedencia de la cópula  $C(u, v)$  y la de cada una de las marginales  $(u, v)$ .

De forma esquemática, los escenarios  $OR^V$  y  $AND^A$  pueden representarse mediante la siguiente figura, donde se reproduce el rango de valores que pueden adoptar las variables  $X$  e  $Y$  y sus respectivos umbrales  $x$  e  $y$ .

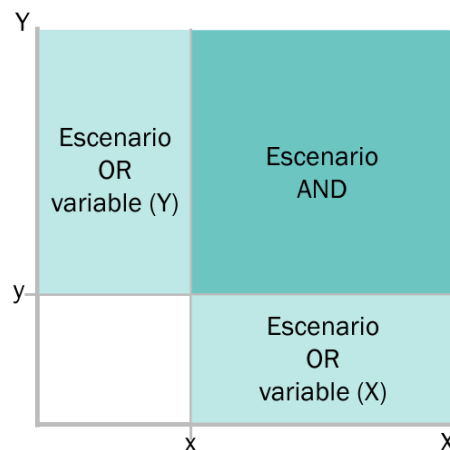


Figura 11 Representación gráfica de los escenarios  $AND$  y  $OR$  para el cálculo del periodo de retorno bivariado

En área en blanco representa la probabilidad de no excedencia de  $x$  e  $y$ , y, por tanto, su cópula  $C(u, v)$ .

El área coloreada por encima, entre la vertical del eje de ordenadas, corresponde con la probabilidad de no excedencia del umbral  $x$ . Por tanto, representa la función de distribución univariada  $u$ .

De forma similar, el área coloreada bajo la horizontal de  $y$  corresponde con la probabilidad de no excedencia de  $y$ , es decir, la función de distribución univariada de  $v$ .

El escenario  $OR^V$ , atendiendo a la fórmula 5, se corresponde entonces con toda el área coloreada, donde se superan el umbral  $x$ , o el umbral  $y$ , o ambos. Es decir, toda el área menos la cópula  $C(u, v)$  en blanco.

El escenario  $AND^A$  (fórmula 6) es entonces el área coloreada en tono más oscuro, donde se exceden ambos umbrales,  $x$  e  $y$ , por tanto, toda el área de la figura 11, menos el área de las distribuciones univariadas  $u$  y  $v$ , más la cópula  $C(u, v)$ , para evitar restarla dos veces.

Los cálculos nos ofrecen una serie de pares de valores para las variables  $X$  e  $Y$  correspondientes a una recurrencia. Representado estos valores se define una curva con dos asíntotas:

- Una vertical, correspondiente con el periodo de retorno univariado de  $X$ .
- Una horizontal, correspondiente con el periodo de retorno univariado de  $Y$ .

Los escenarios  $OR^V$  (figura 12, en azul) y  $AND^A$  (figura 12, en rojo) forman curvas entre las dos asíntotas. Una curva cóncava para el escenario  $AND^A$  y convexa para el escenario  $OR^V$ , compartiendo las asíntotas para cada periodo de retorno dado.

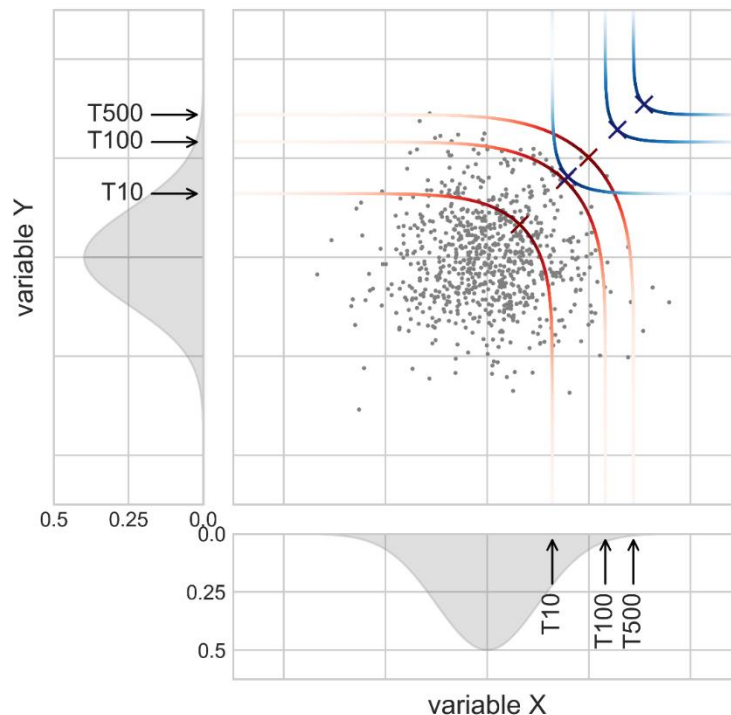


Figura 12 Ejemplo de cópula gaussiana con las curvas de periodo de retorno para los escenarios  $AND$  y  $OR$

En el análisis bivariado, los puntos de las curvas no tienen la misma probabilidad de ocurrencia, aunque se correspondan al mismo periodo de retorno. La cópula





gaussiana representa una campana de Gauss, donde los puntos de mayor probabilidad de excedencia se corresponden con los puntos centrales de la curva.

## 4 Resultados

### 4.1 Análisis univariado

#### 4.1.1 Valores extremos para periodos de retorno predefinidos

Los resultados del ajuste de extremos univariado para la serie de extremos de precipitación, caudal y nivel se muestra en la tabla siguiente. Se incluye la función de distribución con el mejor ajuste, el valor ajustado de los parámetros de dicha función y el valor asociado a cuatro periodos de retorno (25, 50, 100 y 500 años):

Variable	Unidad	Función de distribución				Periodo de retorno			
		Tipo	$\mu$	$\xi$	$\sigma$	25	50	100	500
Precipitación	mm	GEV	59.917	0.204	16.408	98.4	104.0	108.8	117.6
Caudal	m <sup>3</sup> /s	Gumbel	135.69	0	57.718	300	335	369	449
Nivel	m	Gumbel	3.061	0	0.902	5.95	6.58	7.21	8.67

Tabla 6 Ajuste de la función de distribución de extremos y valor extremo para diversos periodos de retorno para la precipitación, el caudal y el nivel

La siguiente figura muestra la función de distribución analítica (línea azul) ajustada para las tres variables anteriores, y la distribución empírica (puntos grises); el punto rojo representa el evento de enero de 2019.

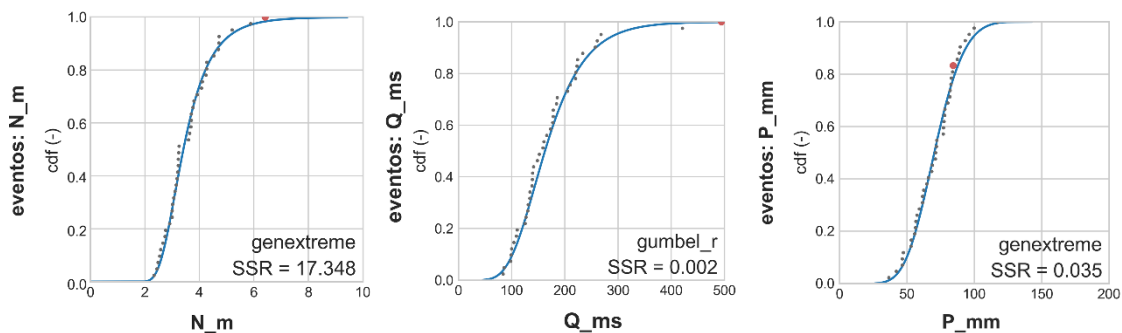


Figura 13 Funciones de distribución empíricas y analíticas para la precipitación, el caudal y el nivel

Mientras que la función ajustada para la precipitación es una GEV, la función con un mejor ajuste para el caudal y el nivel es la Gumbel. En la figura del caudal también puede verse que su posición está en el extremo de la función de distribución de la gráfica.

#### 4.1.2 Periodos de retorno de la inundación de enero de 2019

Para conocer mejor la influencia individual de todas las variables a analizar, se puede extraer su periodo de retorno para evento de enero de 2019 empleando las funciones de distribución anteriormente ajustada. Los resultados para los dos días principales del evento se muestran en la tabla 7; se incluyen los datos para todas

las variables, a excepción de la variable de precipitación del día anterior, por ser los resultados iguales a los de la precipitación diaria.

	Periodos de retorno (años)					
Fecha	Precipitación	Caudal	Nivel	P 2 días	P 3 días	P 6 días
2019-01-23	2	134	27	3	2	2
2019-01-24	5	1200	55	18	20	11

Tabla 7 Periodos de retorno de las principales variables de estudio para los días de inundación

## 4.2 Análisis bivariados

### 4.2.1 Correlación entre variables

Antes de realizar las cópulas, es importante conocer la correlación que existe entre las variables que vamos a analizar, para poder centrar los últimos cálculos en aquellas variables que muestren mejores relaciones.

Estos son los resultados para una correlación de Spearman.

	N_m	Q_ms	P_mm	2d_P_mm	3d_P_mm	6d_P_mm	1d_P_mm
N_m	1.000	0.843	0.263	0.349	0.424	0.549	0.371
Q_ms	0.843	1.000	0.299	0.403	0.474	0.615	0.412
P_mm	0.263	0.299	1.000	0.836	0.721	0.527	0.508
2d_P_mm	0.349	0.403	0.836	1.000	0.888	0.653	0.819
3d_P_mm	0.424	0.474	0.721	0.888	1.000	0.773	0.782
6d_P_mm	0.549	0.615	0.527	0.653	0.773	1.000	0.596
1d_P_mm	0.371	0.412	0.508	0.819	0.782	0.596	1.000

Tabla 8 Correlación de Spearman entre las diferentes variables de estudio

La selección de variables partirá de los mejores valores de correlación obtenidos anteriormente y de los máximos anuales de nivel, caudal y precipitación diaria. Se calcularán por separado, ya que los máximos de unos y otros, en las series de nuestros datos, no siempre coinciden en los mismos días.

Para la precipitación se obtienen los siguientes resultados:

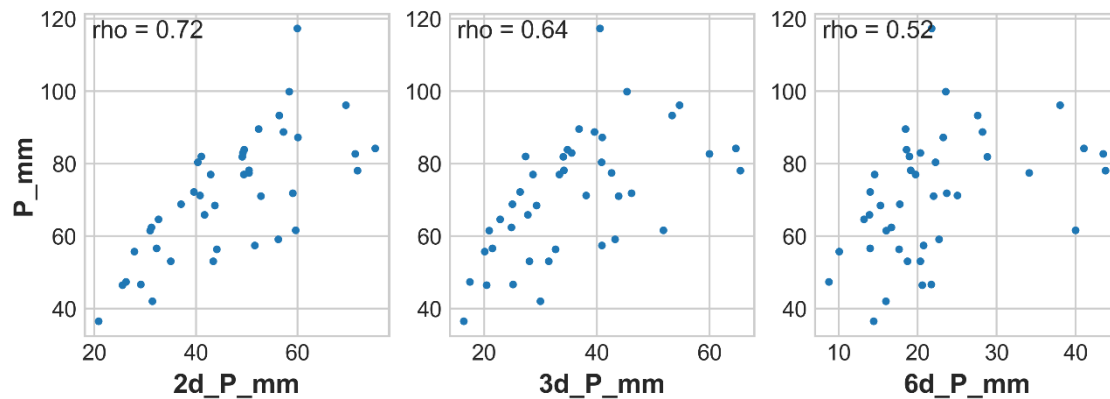


Figura 14 Precipitación máxima y sus principales correlaciones

Según las figuras, en los días de máxima precipitación, la precipitación de los dos, tres y seis últimos días son las variables que mayor correlación tienen. Especialmente la variable de los dos últimos días. El resto de las variables mostraron una correlación menor (Anexo).

Para el caudal:

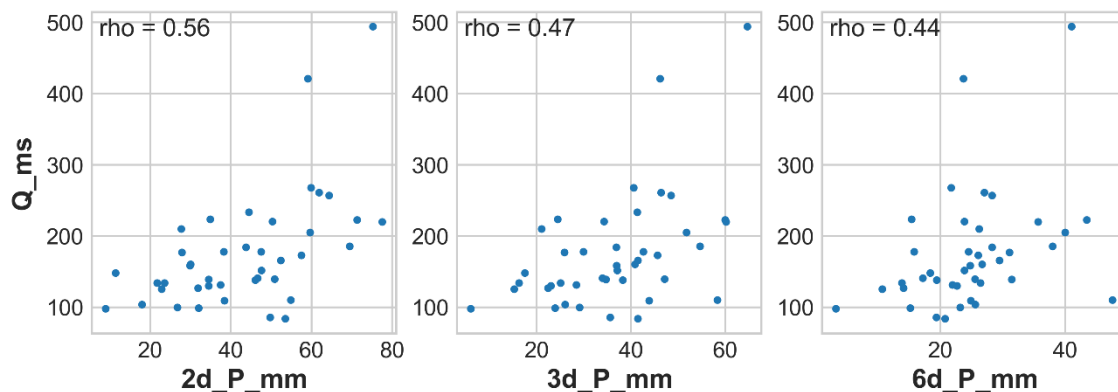


Figura 15 Caudales máximos y sus principales correlaciones

Las figuras muestran las mayores correlaciones entre el caudal máximo anual y las variables de precipitación se encuentra en la precipitación media para los dos últimos días, seguido para tres días y seis días.

Y las correlaciones en los máximos de nivel:

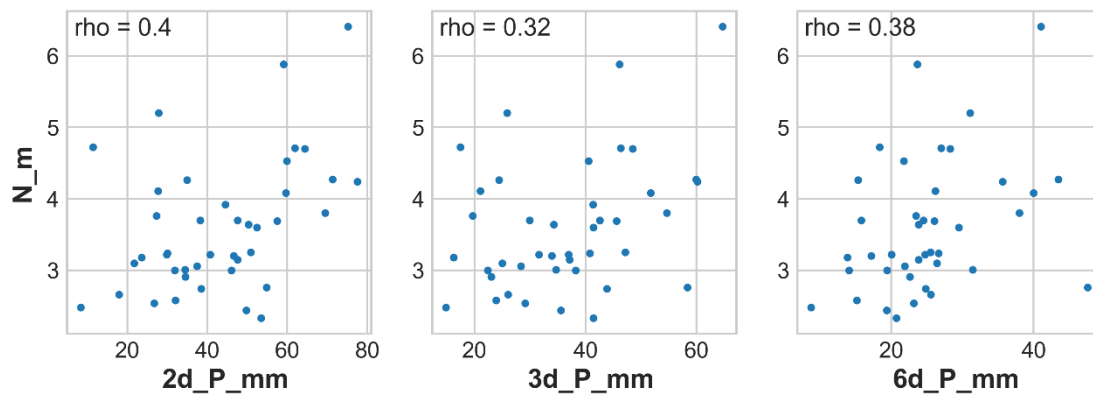


Figura 16 Nivel máximo y sus principales correlaciones

Para el nivel, las figuras muestran una correlación menor que para el caudal, siendo el valor que presenta mejor correlación la media de los últimos dos días.

Observando estos resultados en su conjunto, la variable de dos días es la que tiene una mayor correlación con el resto de los valores. Bajo estas correlaciones se harán las cópulas usando el caudal y nivel máximos como variable a analizar y las variables de precipitación para analizar con las cópulas.

#### 4.2.2 Cópulas gaussianas

En función a los resultados obtenidos en los puntos anteriores, se decide emplear como variable objetivo a analizar en las cópulas el nivel máximo anual del río y las variables que se emplearan en las cópulas las asociadas a estos valores extremos de precipitación, intensidad media a dos días e intensidad media a tres días.

Para las variables de precipitación en el día y media de los dos últimos días, obtenemos la siguiente figura representando la cópula gaussiana obtenida y sus periodos de retorno.

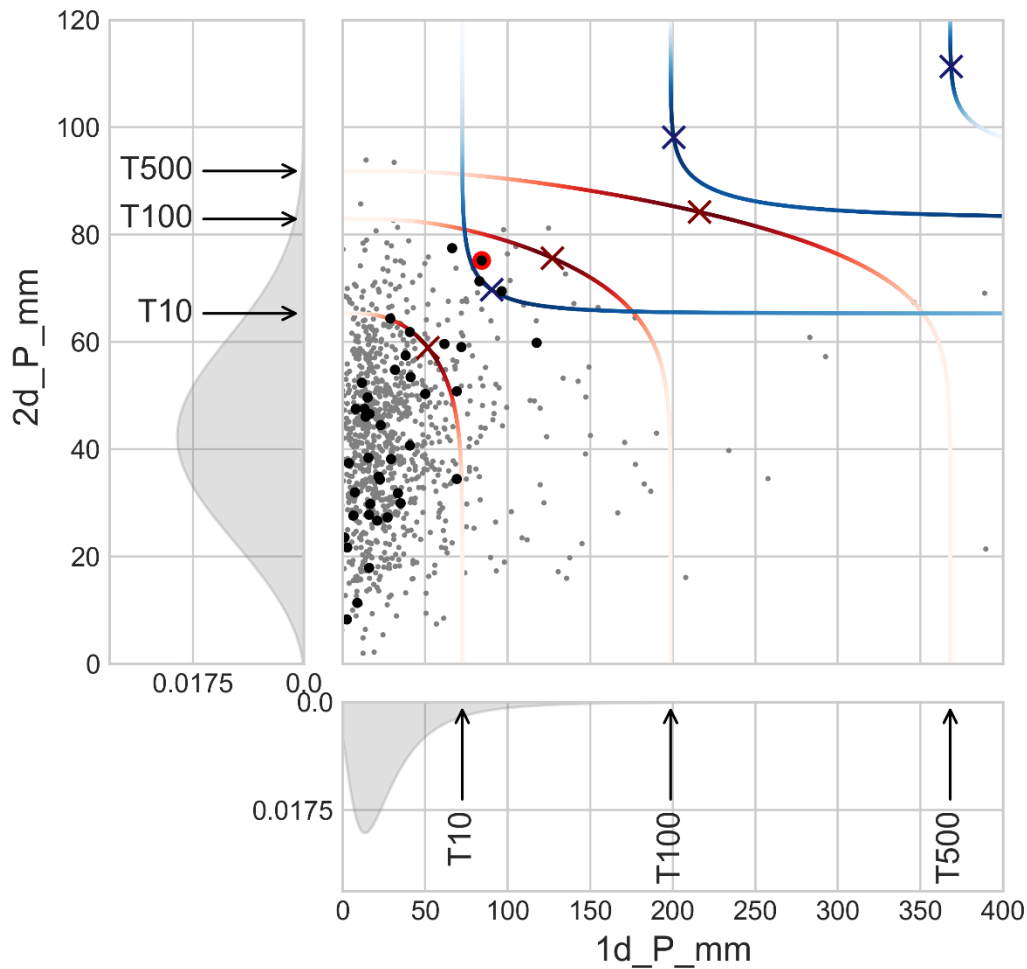


Figura 17 Representación de la cópula entre las variables de precipitación y media de la precipitación de dos últimos días.

La figura representa con los puntos en negro los valores observados de precipitación diaria e intensidad de los dos días para los valores máximos de nivel de la cuenca. Los puntos grises son valores aleatorios obtenidos aplicando nuestro modelo de cópula. Las líneas rojas representan los escenarios AND para los periodos de retorno de 10, 100 y 500 años; marcados también en los laterales con las campanas de Gauss para cada variable. La línea azul representa el escenario OR, para los mismos periodos de retorno. La intensidad de colores en las líneas señala las zonas de mayor probabilidad de ocurrencia de un evento. El punto marcado en rojo corresponde al evento del 24 de enero del 2019.

Los puntos de mayor probabilidad representados en los vértices de las curvas de las gráficas serían los siguientes.



	Escenario AND		Escenario OR	
	2d_P_mm	1d_P_mm	2d_P_mm	1d_P_mm
<b>10</b>	51.6517	58.9213	90.0901	69.7242
<b>100</b>	126.927	75.6461	200.544	98.1381
<b>500</b>	216.117	84.2042	368.605	111.351

Tabla 9 Tabla con los valores de los puntos de mayor probabilidad para cada periodo de retorno para las variables de precipitación y precipitación media de dos días

Se realizan las mismas operaciones empleando las variables de precipitación a dos días y a tres días, con una mejor correlación con los máximos de nivel. Obteniendo:

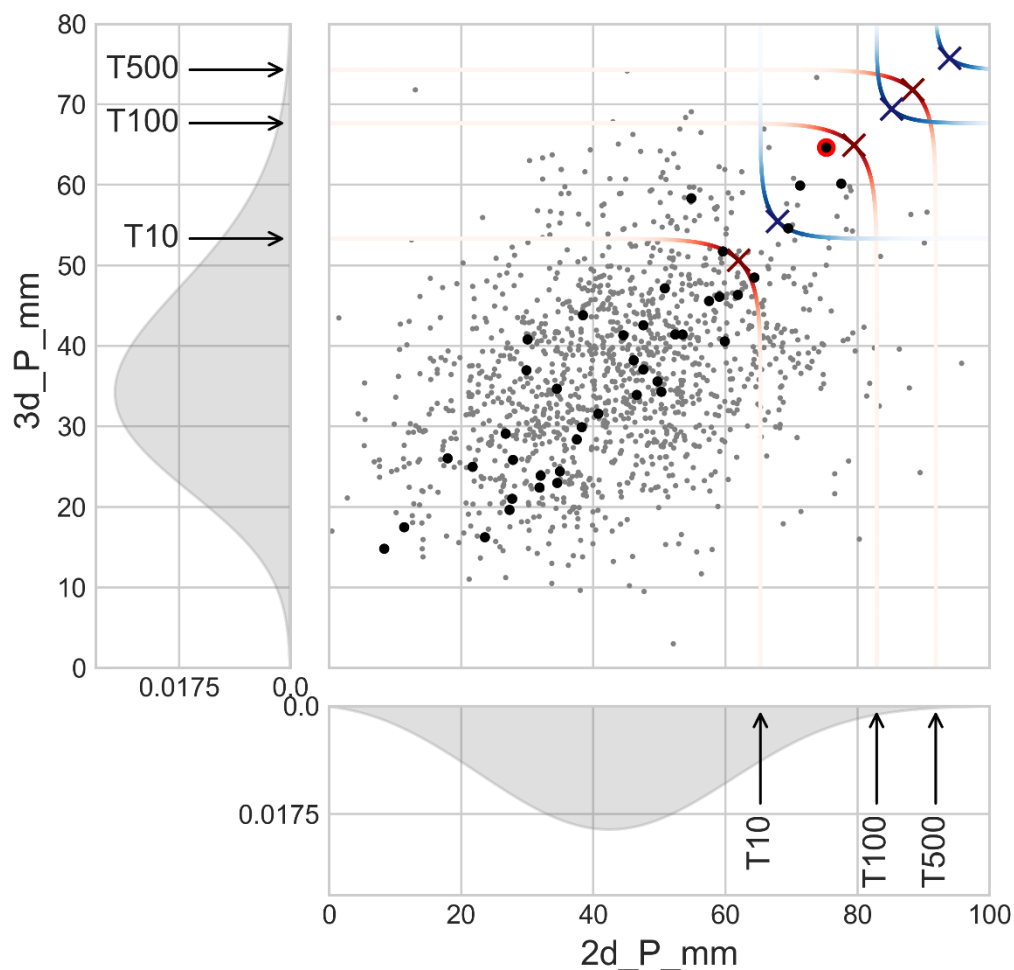


Figura 18 Representación de la cópula entre las variables de media de la precipitación de dos y tres últimos días, con sus probabilidades.

Esta figura representa la distribución de los valores para las variables de intensidad de lluvia para los dos y los tres últimos días, para los días de máximo nivel anual. Como en la figura anterior, las líneas rojas representan los periodos de retorno (a 10, 100 y 500 años) para el escenario AND y para el escenario OR.



	Escenario AND		Escenario OR	
	2d_P_mm	3d_P_mm	2d_P_mm	1d_P_mm
<b>10</b>	61.962	50.6575	67.8756	55.4955
<b>100</b>	79.4497	64.9449	85.1349	69.4294
<b>500</b>	88.3181	71.8318	93.8939	75.7458

*Tabla 10 Tabla con los valores de los puntos de mayor probabilidad para cada periodo de retorno para las variables de precipitación media de dos y tres días*

## 5 Discusión

En el análisis univariado destaca que, a excepción del caudal, ninguna de las variables llega a valores de periodo de retorno de 100 años, si bien nivel se acerca. Lo cual podría considerarse que los datos no representan correctamente la inundación, o que la inundación ha ocurrido pese a que no son valores que deberían desbordar los sistemas de gestión y contención de inundaciones.

El caso del caudal, se obtiene un periodo de retorno de 1000 años, que parece poco plausible, la explicación podría encontrarse en el dato para el día 24 de enero, de casi 500 m<sup>3</sup>/s, es el máximo histórico la serie, quedando a la cola de la distribución (Figura 13) por lo que el periodo de retorno puede ser sobrevalorado.

En el caso de las variables de la precipitación, puede apreciarse que ninguno representa un valor extremo, individualmente solo el valor de la precipitación media de los últimos dos y tres días parecen relevantes, al encontrarse alrededor de los 20 años de periodo de retorno, que se acerca más a los 55 años del nivel. Los valores de precipitación para el día 23 y el día 24 apenas llegan a un periodo de retorno de 2 y 5 años. respectivamente. Ninguna de estas variables individualmente parece ser suficiente para explicar ni el valor del nivel en los datos, ni la magnitud de la inundación.

Asumiendo que los datos de nivel y caudal que tenemos son correctos, la severidad de la inundación superó los sistemas de gestión, habitualmente diseñados para inundaciones de 100 años de retorno, con unos periodos de retorno sensiblemente inferiores, podría deberse a los siguientes escenarios:

- La metodología utilizada en la definición de los periodos de retorno de avenida es inadecuada.
- Las infraestructuras y medidas de gestión de inundaciones aplicadas en la cuenca están perdiendo su efectividad (lo que podría explicar también inundaciones en la cuenca tan recientes como 2016).
- El comportamiento hidrológico de la cuenca está cambiando en los últimos años por alguna razón y los cálculos realizados en su momento ya no sirven para la situación actual.

Este informe analiza el primero de estos puntos, la idoneidad de la metodología habitualmente empleada. En primer lugar, los resultados muestran que la precipitación acumulada en dos días influye en mucha mayor medida en la magnitud de la inundación que la precipitación diaria, incluso la precipitación media de tres y seis días parece tener mayor influencia en avenidas extremas que la diaria.

En segundo lugar, estas variables por sí solas no permiten explicar la magnitud de la inundación de 2019.

Estas conclusiones confirman la suposición planteada al inicio del análisis de que el valor de la precipitación en el día de las inundaciones es insuficiente para explicar la gravedad de estas.

Para analizar cuáles podrían ser las variables que expliquen esta inundación, se realiza el análisis bivariado.

Puede observarse, analizando las relaciones entre las variables, es que, con mucha frecuencia, los máximos de una variable para un día no coinciden con los valores máximos de las otras variables para ese mismo día, como se podría esperar. Se puede ver en las gráficas como la Figura 15 que, para intensidades de precipitación altas, en muchos días los valores de nivel y caudal son bajos, y viceversa.

Estas aparentes discrepancias quizá podrían explicarse incluyendo otras valoraciones como la estación del año y temperatura ambiente, pero incluirlas añadirían demasiada complejidad para el alcance de este análisis.

En el cálculo de las cópulas, puede observarse que la figura 17 no parece representar correctamente los valores de periodos de retorno, deformándose demasiado las curvas, especialmente las del escenario OR. Esto podría deberse a que aunque exista una buena correlación entre las variables de precipitación diaria y de media de dos últimos días, los valores de la distribución de la lluvia diaria, para extremos de nivel y caudal, presentan grandes diferencias que, al configurar el modelo de distribución de la cópula, hacen que las probabilidades generadas se extiendan muy por encima de lo que pueden abarcar el periodo de retorno de las otras variables.

De todas formas, el punto para el día 24 de enero sí está representado en la gráfica y se puede apreciar que en el escenario AND queda cerca del periodo de retorno de 100 años, por lo que se puede deducir que la ocurrencia de ambas variables ofrece un periodo de retorno más cercano a lo ocurrido en los días de la inundación que los valores individuales.

Las cópulas para las variables de intensidad media para los dos y tres últimos días, para los valores máximos de nivel, ofrecen una gráfica más legible. Observado la figura 18, puede verse que el día de la inundación no solo corresponde a un valor que se acerca a un periodo de retorno de 100 años en el escenario AND, también que se encuentra cerca de lo que el modelo señala como área más probable para este valor.

De este análisis de las cópulas se podría concluir que la probabilidad de que un evento extremo ocurra es mayor al considerar la ocurrencia de dos variables al mismo tiempo, sin que estas necesariamente representen un valor extremo por sí mismas. Es decir, los valores de precipitación para uno, dos o tres días, no necesitan ser valores inusuales para que generen un evento de inundación, y la probabilidad de que estos valores recurrentes ocurran al mismo tiempo es mayor que la probabilidad de que una única variable genere una inundación, como suele asumirse en los cálculos. Siendo esto así, el análisis de extremos habitualmente utilizado en el diseño de estructuras hidráulicas queda invalidado, puesto que al realizarse sobre una única variable infravalora el riesgo.

Por el momento, estas observaciones solo serían adecuadas para el caso estudiado, que es aplicado para un solo evento en una cuenca pequeña, y ya que, observando las correlaciones entre las variables, aún hay bastante irregularidad entre las mismas. Sería interesante incluir otras variables que no se están teniendo en cuenta en este informe y que pueden variar algunas de las relaciones entre las



variables, influyendo en la generación de escorrentía, aportando información extra en la generación de la cópula (por ejemplo, la nieve acumulada en la cuenca). Este estudio se ha limitado a cópulas gaussianas bivariadas; que puede variar algunas de las relaciones entre las variables y ofrecer información más precisa para los cálculos al permitir incrementar el número de dimensiones, es decir, el número de variables predictoras.

## 6 Propuestas

Partiendo de que la conclusión de que el cálculo de periodo de retorno empleando una sola variable es insuficiente, sería importante crear nuevas metodologías para calcular los periodos de retorno empleados en el diseño de infraestructuras y medidas de mitigación de inundaciones que se adapten mejor a las diferentes cuencas.

Para el caso de las infraestructuras y sistemas de mitigación ya construidos, sería necesaria una reevaluación para analizar que cumplen con los parámetros de seguridad para los que fueron construidos y que no han sufrido un deterioro significativo que haga que ya no sirvan a su propósito inicial.

Para la cuenca que nos ocupa, considerando que los periodos de retorno obtenidos siguen siendo en apariencia insuficientes para la gravedad de las inundaciones ocurridas, se podría hacer un análisis en más detalle de la cuenca. Un análisis de tendencias podría ofrecer información sobre si existen variaciones en el comportamiento hidrológico del río que expliquen con más profundidad la serie de recientes inundaciones.



## 7 Referencias

Gobierno de Cantabria (2010). “Decreto 57/2010, de 16 de septiembre, por el que se aprueba el Plan Especial de Protección Civil de la Comunidad Autónoma de Cantabria ante el Riesgo de Inundaciones, INUNCANT”. BOC.

Manuela Irene Brunner, Jan Seibert and Anne-Catherine Favre (2016). “Bivariate return periods and their importance for flood peak and volume estimation.” WIREs Water 2016, 3:819–833. doi: 10.1002/wat2.1173.

A. I. Requena, L. Mediero, and L. Garrote (2013). “A bivariate return period based on copulas for hydrologic dam design: accounting for reservoir routing in risk estimation.” Hydrol. Earth Syst. Sci., 17, 3023–3038, 2013. doi:10.5194/hess-17-3023-201

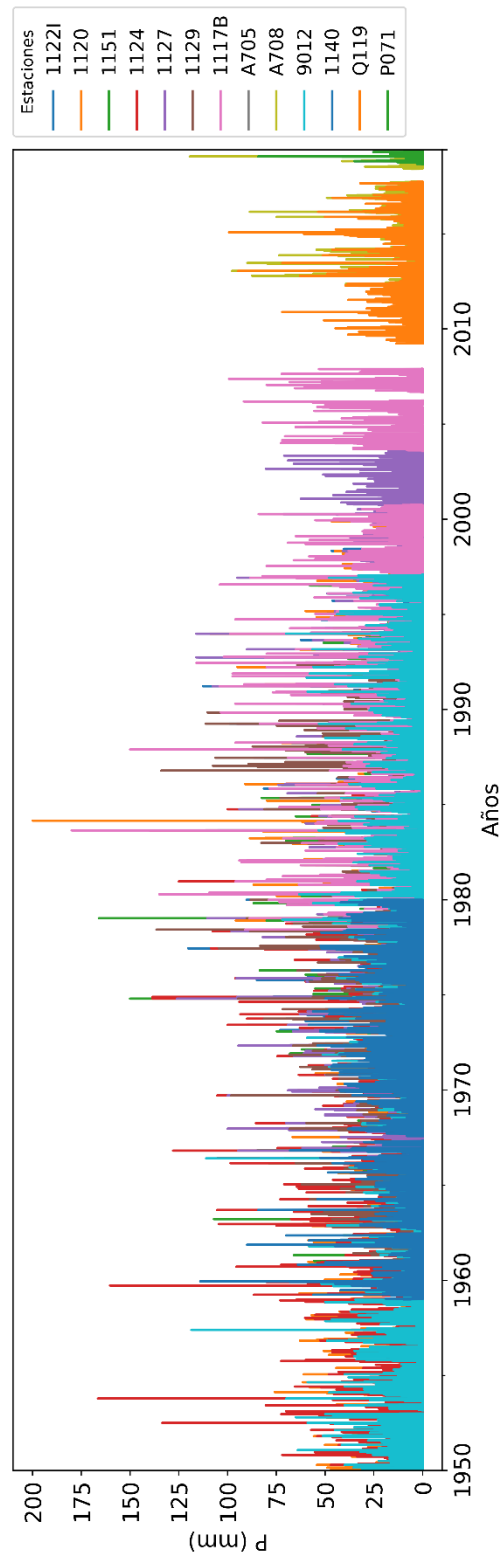
Christian Genest and Anne-Catherine Favre (2007). “Everything You Always Wanted to Know about Copula Modeling but Were Afraid to Ask.” 10.1061/(ASCE)1084-0699(2007)12:4(347)

Ben Alaya, M. A., Chebana, F., and Ouarda, T. B. M. J. (2014). “Probabilistic Gaussian Copula Regression Model for Multisite and Multivariable Downscaling.” Journal of Climate, 27(9), 3331–3347.

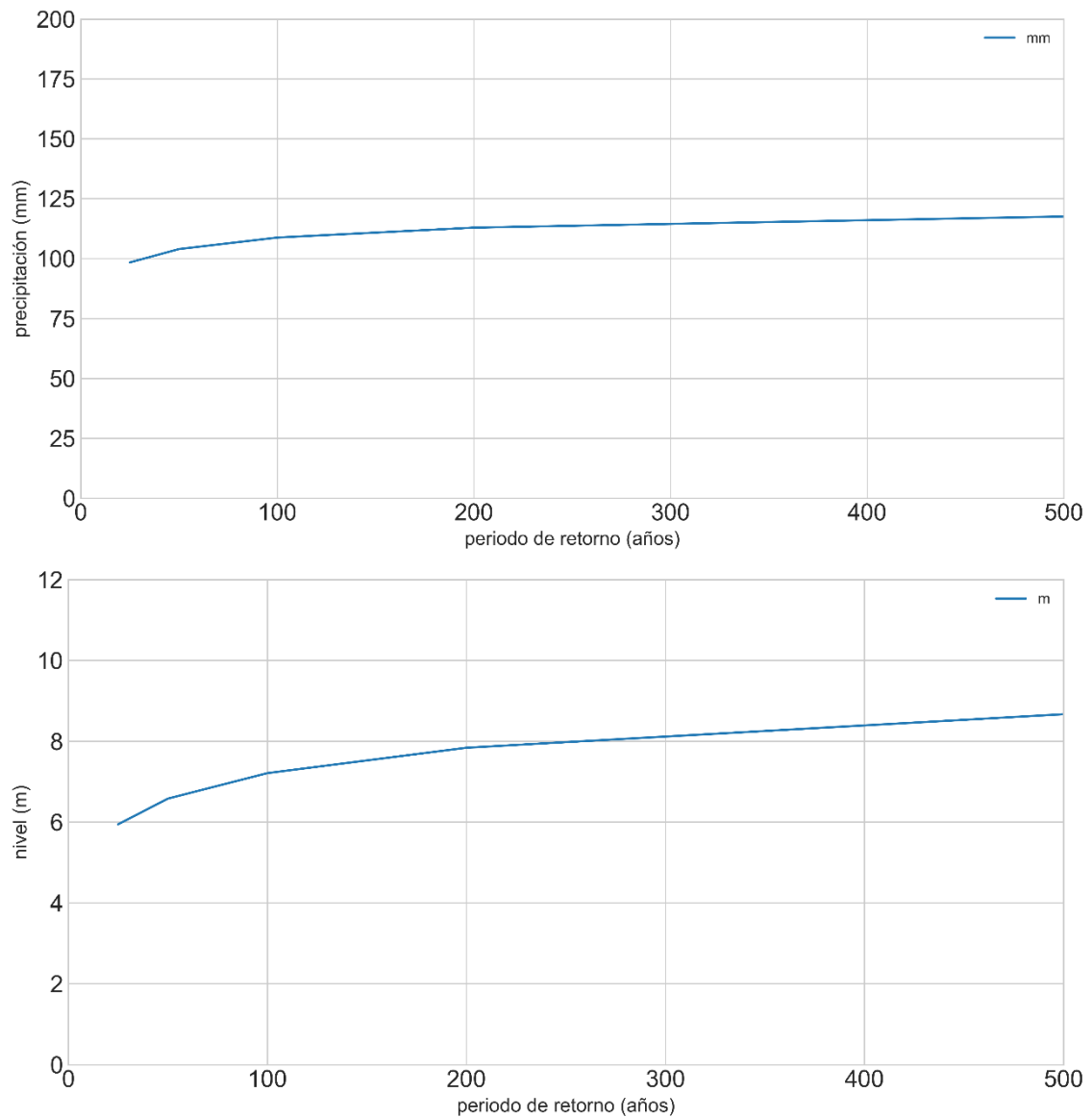
Ben Alaya, M. A., Chebana, F., and Ouarda, T. B. M. J. (2015). “Probabilistic Multisite Statistical Downscaling for Daily Precipitation Using a Bernoulli–Generalized Pareto Multivariate Autoregressive Model.” Journal of Climate, 28(6), 2349–2364.

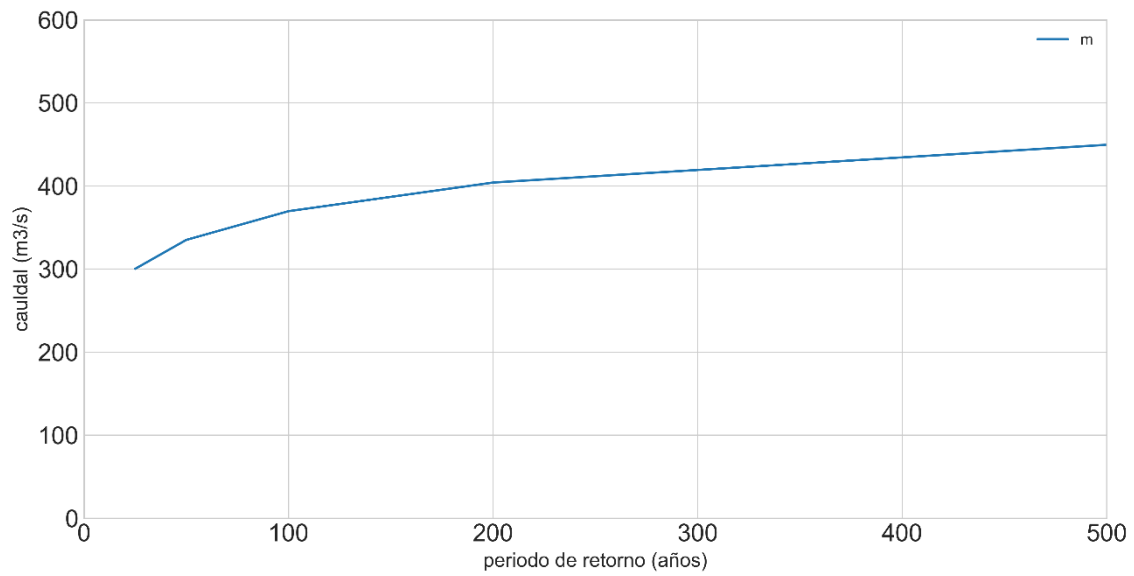
## 8 Anexos

### 8.1 Datos de precipitación de las estaciones seleccionadas para la cuenca Ao88 entre 1950 y 2019

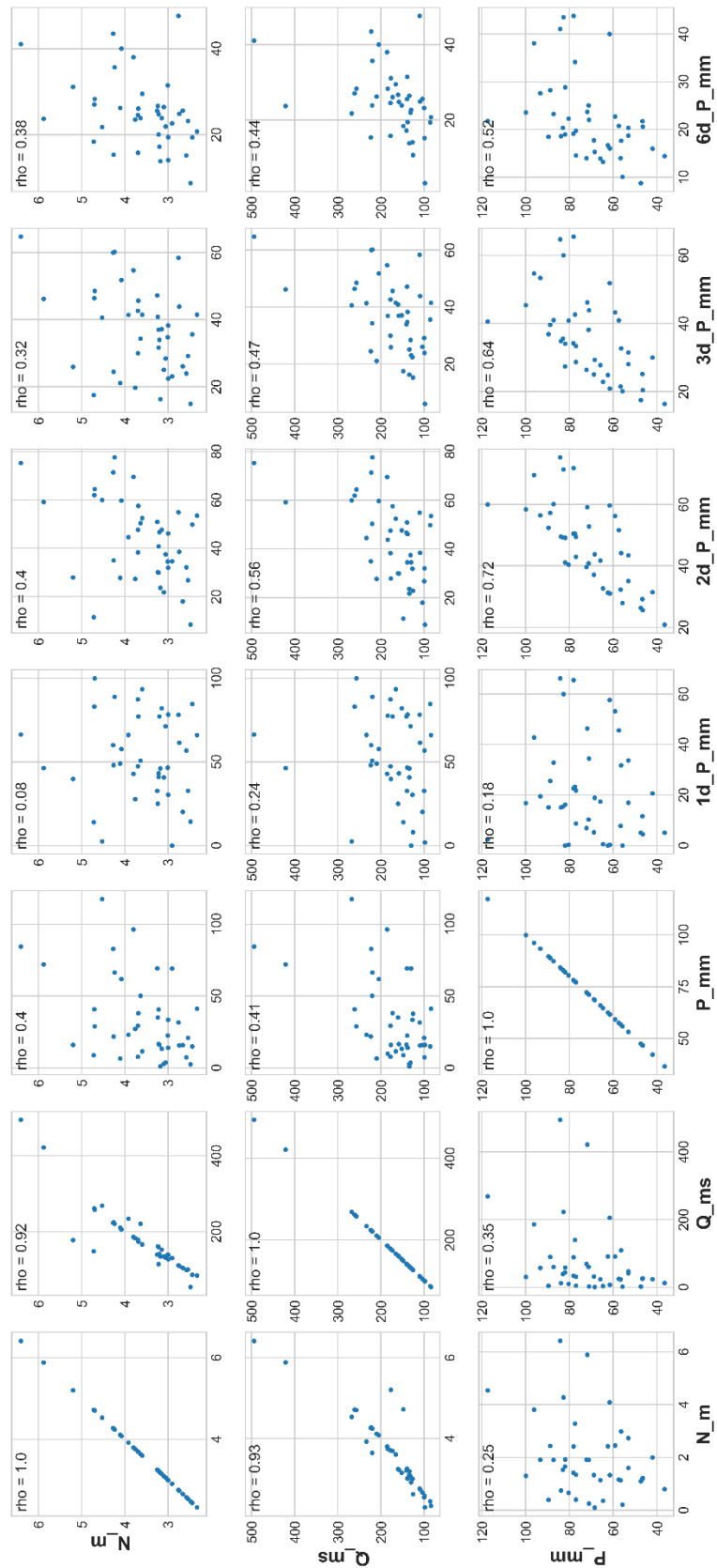


## 8.2 Periodos de retorno para los valores de precipitación, nivel y caudal





### 8.3 Correlación entre los extremos



## 8.4 Ajuste de la distribución de extremos

